AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

UNIVERSITY OF CALIFORNIA
DAVIS
SEP 6 1977
Agricultural Economics Library

August 1977 Study Paper #77-6

A Review of Old and New Methodology for Distribution Research

by

Don C. Reading and John E. Keith

Contributed paper presented at AAEA-WAEA joint meetings, San Diego, California, July 31-August 3, 1977.

ABSTRACT

A Review of Old and New Methodology

For Distribution Research

Attemps to measure distribution changes have fallen into categories: measures of personal income or wealth distribution called personal distribution and measures of returns to factors of production called functional distribution. Specific measures of personal distribution consist of the traditional formula measures (Gini, Pareta, etc.) and the more recent use of functional measures or estimations (Beta, Gama, etc.) Functional distribution research utilizes production functions and other output models to measure changes either quantiles of income or in formula or functional measures.

Evaluation of the various distribution measures depends upon the requirements of the specific research task. Care must be taken to temper the analysis with non-operational structural impacts. Future problems will be the "old" problems, such as the definition of inequality, as they relate to new areas of research.

A REVIEW OF OLD AND NEW METHODOLOGY FOR DISTRIBUTION RESEARCH

By Don C. Reading and John E. Keith*

INTRODUCTION[1/] (4,5,33)

The distribution of income (a flow) and wealth (a stock) among members of a society is a function of the system of exchange (the market), government policy, and the interaction of the two components. Income and wealth are accumulated as a result of ownership of resources, and the ability to utilize owned resources to capture returns. Clearly, property rights and public policy can determine, or at least affect changes in, the distribution of income and the distribution of wealth as well as relationships between the two.

Natural resource policy can have both direct and indirect effects on distribution. Direct income transfers between individuals are most often associated with other kinds of policies, such as food stamps or welfare payments, but fees and charges for natural resource use, particularly if those charges discriminate among users in some way, can cause income transfers. Indirect effects can be generated by restricting the use of resources or the availability of the resource. Owners of substitute resources can reap the gains to scarcity in the form of higher wages and prices, while users of the resource must contribute greater portions of their wages or wealth in order to satisfy their demand. Owners of resources which are restricted suffer losses of returns in a similar manner.

---

*Associate Professor of Economics, Idaho State University and Assistant Professor of Agricultural Economics, Utah State University.

[1]No citations are included in the text of this paper. Instead a bibliography is included and numbered. The numbers of related references are listed at the beginning of each section.

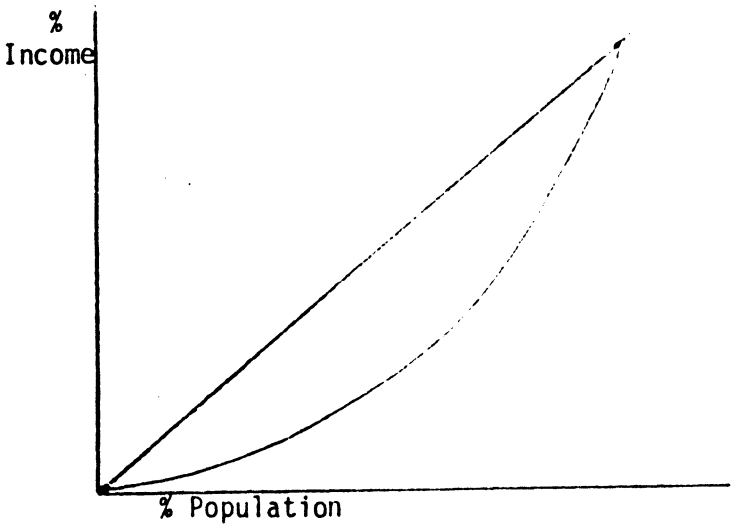

Figure 1. Lorenz Curve

the whole distribution. A comparison of an index in time series or cross sections will yield information about how the distribution has changed.[2/] There are several of these formula measures, the most commonly used of which are discussed below:[3/]

The Gini Coefficients is probably the most widely known of the formula measures. This coefficient is a measure of the area between the Lorenz curve and the 45^{o} line. The Gini can be mathematically defined as

$$G = (1/2\ n^2\mu)\ \Sigma_{i=1}^{n}\ \Sigma_{j=1}^{n}\ |y_i - y_j|$$

where n = number of individuals

μ = mean income

y_{ij} = income to person i or j

[2/] It is interesting to note that most efficiency studies assume a given income distribution before and after the analysis of policy effects, yet never tests for changes in the distribution. Clearly these efficiency analyses are flawed since there is a probability of a distributional charge, as Samuelson and others have pointed out.

[3/] Much of this discussion can be found in Amartya Sen's book, On Economic Inequality.

Changes in policy, then, may have significant effects on the distribution income. Whether these changes are "good" or "bad" is a matter of debate; the impact should at least be analyzed.

Attempts to measure changes in income distribution have fallen into two categories: measures of personal income or wealth distribution called personal distribution, and measures of returns to factors of production called functional distribution. As has been pointed out, these two categories are linked through the pattern of ownership of productive factors, and some recent studies have attempted to determine procedures to identify this relationship. Since personal income distribution is the main topic of this symposium, the paper will focus primarily on the methodology of measuring personal distribution, and the attempts to relate functional distribution to personal distribution.

PERSONAL DISTRIBUTION (9, 14, 34, 36)

The measurement of personal distribution has a relatively long history in economic literature. The most used and oldest methods of measurement are "Formula Measures", but recently there has been considerable interest in other approaches, mainly in "Functional Measures". Basic to all of these measures is the Lorenz Curve, which relates percentage of income or wealth of the total population to the percentage of the total population which holds that income or wealth. This curve is illustrated in Figure 1 and measures the deviation of the actual distribution from equal distribution (a 45 degree line).

Formula Measures (1, 2, 6, 8, 10, 15, 16, 21, 22, 28, 30)

The formula measures simply use given quantiles of income distribution - that is, the populations which fall within discrete categories of income, earnings or wealth - and attempt to find an index or indication which represents

The Pareto measure is also one of the oldest measures used. This measure describes the number of individuals with income over a given amount. This formulation can also be used to estimate the midpoints of the unbounded upper quantile of discrete income data. (U.S. Bureau of Census).

The Theil entropy measure is the average of the logarithms of the reciprocals of income shares weighted by the income shares of each individual. Its mathematical formulation is:

$$T = \Sigma_{i=1}^{n} X_i \log nX_i$$

where X_i is the relative share of income going to each person.

The general statistical measures include measures of both central tendency and of dispersion. There are several of these measures. Relative mean income is a measure of the mean income for each of the income quantiles compared with the mean for the total population. Changes in this measure over time indicates that a given quantile is gaining or losing relative to the total population.

Relative mean deviation is a measure of the deviation of each individual's income from the mean income relative to total income. The measure expressed mathematically by:

$$M = \Sigma_{i=1}^{n} |\mu - y_i| / n\mu$$

Variance and coefficient of variation measures use the deviations from the mean also. Mathematically variation is:

$$V = \Sigma_{i=1}^{n} (\mu - y_i)^2$$

The coefficient of variation is simply the deviation divided by the mean:

$$C = \sqrt{V}/\mu$$

The standard deviation of logarithms measure accentuates the deviations in the lower income groups more than do the absolute value measures.

Mathematically:

$$\text{II} = \{\Sigma_{i=1}^{n} (\log \mu - \log y_i)^2/n\}^{1/2}$$

Measures of skewness have also been suggested as distribution measures. A major problem with skewness measures is that inequality changes can occur with symmetric distributions as well as non-symmetric.

There have been other measures suggested and used in the measurement of personal income distribution. The Pietra index is a measure of the area between the Lorenz and equal distribution curves which uses a geometric methcd of fitting triangles within that area as an approximation. Kurtoses measures have also been suggested.

There are other measures which are based not on the objective quantification of distribution but are based on normative criteria. Those are the Dalton measure and the Atkinson Measure. Mathematically they may be expressed as

$$D = \{\Sigma_{i=1}^{n} U(y_i) / nU(\mu), \qquad \text{and}$$

$$A = 1 - (\{\Sigma_{i=1}^{n} U(Y_i)\}/\mu).$$

Where $U(Y_i)$ is in some sense the welfare generated from given levels of income.[4/] For a given assumption about the functional form of the social welfare function, these measures reduce to more commonly used objective measures such as the Gini, the Theil, the coefficient of variance, or the standard deviation of logrithms. It could probably be shown that each of the other formula measures would correspond to the Dalton or Atkinson measure, given alternative forms of the welfare function.

[4/] Note that these are <u>not</u> utility functions, but relate directly to social welfare function.

All of these measures have weaknesses and strengths, depending upon the purpose for which they are used. However, all of them also have a common weakness: each uses a single measure to describe the actual income distribution or the deviation of actual from the equal distribution line. The implication is that the same index number can be generated by many (in fact, infinitely many) Lorenz Curves. Thus, these indices are not unique. Where Lorenz Curves cross, but yield the same formula index, personal income distribution may be either more or less equal with respect to high and low income quantiles relative to the equal distribution line. This ambiquity has led researchers to look for unique measures of distribution in the various families of probability density functions.

Functional Measures (11, 12, 22, 24, 25, 26, 31)

The advantage of functional measures are that more than one parameter or variable can be used to identify the distribution. These approaches use distribution functions to estimate or approximate the Lorenz curve itself, rather than the area between the curve the equal distribution line, the relative mean, or the dispersion characteristics. There are several probability density functions which have been suggested for estimating distributions. However, a given probability density function which is the statistically "best" fit for a specific Lorenz curve may not yield the best fit for other income distributions. If regional, racial, occupational, or other classifications are desired, a determination of the "best" fit must first be made.

The general approach in using these functions has been to estimate the parameters of the density functions for the observed income distribution. The midpoints of each discrete income quantiles have been used as the observations, with the midpoint of unbounded highest income quanitle estimated by

Pareto-Levy curve or similar functional form. Some of the more commonly suggested density functions are discussed below.

Logarithm functions have been used. The lognormal and the displaced lognormal functions have been examined by Metcalf, and others. The displaced lognormal function is the more appropriate function since the lognormal is a special case of the displaced lognormal. The mathematical formula is:

$$L = \frac{1}{(X-C)\beta\sqrt{2\,\Pi}} \quad \exp - \frac{(\log (X-C - \log \alpha)^{2}}{2\beta^{2}}$$

Where X is the income, α and β are parameters, and C is the skewness variable. Note that when C = 0, L is the lognormal distribution. These functional forms have proved to somewhat inadequate for estimating national Lorenz curves.

Two functional forms of the Pearson distribution family have been tested relatively widely: the gamma and the beta functions. Both have been found to be "better" estimations of the Lorenz curve than the log functions. The mathematical formulae are:

$$G = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \quad X^{\alpha-1} e^{-\beta X} \underline{5/}$$

where X is the income variable; and α and β are the parameters; and

$$B = \frac{\Gamma(\alpha,\beta)}{\Gamma(\alpha)\Gamma(\beta)} X^{\alpha-1}(1-X)^{\beta-1}$$

where X is an index of incomes ($0 \leq X \leq 1$), and α and β are the parameters.

It has been shown that the gamma function is a special case of the beta, where α approaches infinity. It has also been shown that the Gini, Theil, and other single-valued indices are a function of the first parameter, α, of the gamma density. Thus, the non-uniqueness of the single parameter function is clear.

$\underline{5/}$ $\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy$ Where $y = (\frac{X}{\beta})$

In studies comparing the two functions for national and SMSA data, the beta function appears to be the better estimator of the Lorenz curve, although for other selected populations this conclusion may not hold.

Other distributions have been suggested by several authors, including the Weibul, the Pareto, the SECH, and the Champernowne distributions. Some of these distributions are, as in the case of the gamma and beta, special cases of more general distributions. Each of these distributions has been shown in specific cases to be relatively good estimators, although a full comparison of all the functions over several different populations has not been attempted.

While the functional-form estimators of the Lorenz curve appear to be more appropriate than the single parameter measures, the statistical analyses which can be performed are limited to estimations of the parameters of the function, in an ex post sense. In order to utilize these estimators in policy analysis, the relationships between the policy and the Lorenz curve must be conceptualized. The second class of distribution, functional distribution, is one way of approaching the analysis of policy effects.

FUNCTIONAL DISTRIBUTION

Economics literature abounds with applications of, and theoretical additions to, functional distribution in the form of marginal productivity theory and factor shares analyses. It is not the purpose of this paper to explore these topics, except to the extent that functional distribution is related to personal distribution through ownership of, and associated property rights to, the factors of production. The owners of the factors will extract the rents; the relative position of those who own high or low rent-earning factors will determine the distribution of income, or of wealth.

Factor Shares and Production Functions (3,7,18,19,20,22, 27, 29, 32, 39, 40)

A relatively new thrust in income distribution research is to conceptualize and test ways in which in the functional distribution is translated into personal distribution. A wide variety of approaches have been used to generate these linkages. One basic methodology has been to estimate the activity or output in a given industrial sector, and to relate the incomes of factors employed by that sector to person distribution. Labor has been the principle factor analyzed, but some studies have utilized capital ownership as well. Data are used to translate changes in the returns, or value added, in a sector into increases in income by quantile. In this way, policies which generate different changes among sectors can be studies for their personal income distribution effect. Figure 2 is a schematic of these procedures.

The economic output model used has varied. Input-output tables are a common approach, wherein the direct and indirect effects of policy on all sectors is examined. Mathematical programming, simulation and general equilibrium models have also been used as economic models. There is no reason to exclude interfacing of any or all of the economic models to analyze the impact of policy on the economic sectors, although such combinations are relatively uncommon in income distribution research.

Factor employment and earned income, by sector, has also been estimated using various techniques. Employment by skill level or occupational title by industry has been used, as have coefficients of total labor and capital factor shares derived from Cobb-Douglas production functions. The primary problem with the Cobb-Douglas approach has been the distribution of returns to capital by income quantile. A few attempts have been made to incorporate

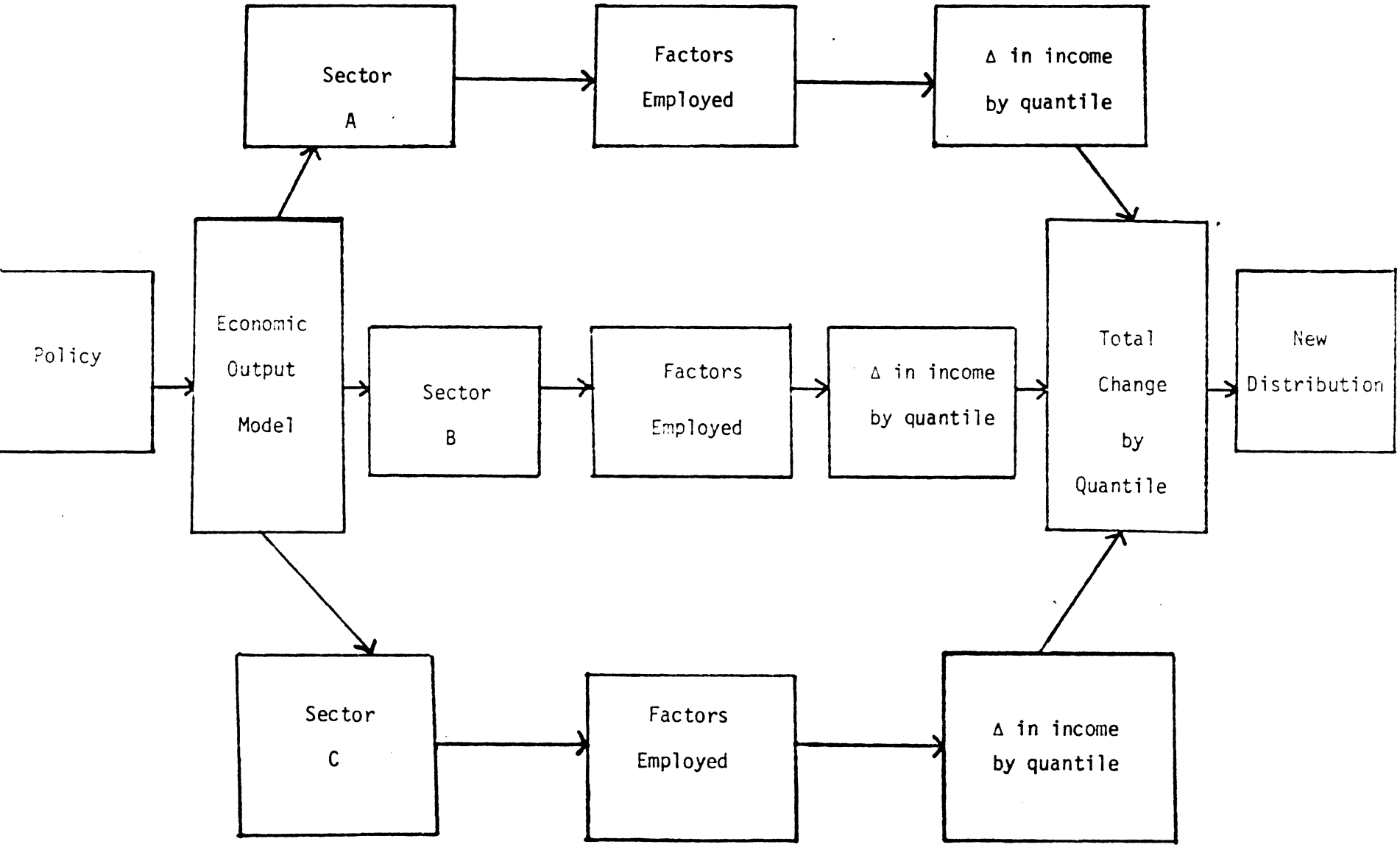

Figure 2. Policy Analysis

returns to owned capital by using averages of reported capital income by occupation. Given the assumptions about capital income, the fraction of total income to a given sector which is paid to each skill level or occupational type is obtained. The average income of each skill level or occupational type is used to establish the income quantiles in which the individual belongs. Sectoral changes are assumed to generate proportional change in its associated quantiles. The sum of all sectoral impacts yields the new income distribution by quantiles. The direct use of quantile changes for distribution impact estimations has been termed a graphical analysis. Some information is lost when these quantiles are highly aggregated, and a direct causal relationship between policy or other induced change in a sector is not estimated.

Other approaches, also based on productivity and market equilibrium, generally derive structural equations for given sectors and/or factors of production. Partial and general equilibrium models have been employed in these efforts. The structural equations are in turn utilized in econometric models which estimate either the income distribution indices (formula measures) or the parameters of Lorenz curve estimators (functional measures) directly from existing data. The regression coefficients indicate the impacts of policy changes. Estimations of income classes and employment categories by sector or industry have also been used in econometric approaches, either to break down sectoral distribution or to generate structural equations.

The production-function-based approaches ignore those institutional constraints which play a role in the distribution of income, except when these constraints are explicit in the model. These institutional constraints

may be difficult to quantify, but may be critical when attempting to assess the impacts of policy.

Structural Models (13, 32, 38)

The term "structural model" is used to denote applications of models which explicitly consider the institutional impacts on income distribution. Friedman's now classic 1953 article would be an example of this approach. In it he states that the distribution of income is a function of nonpecuniary factors as well as society's risk preference. While, a priori, it is rational to assume these institutional factors are important in the study of income and wealth distribution, it is difficult to operationalize them in a meaningful way. At a minimum however, they should be considered, particularly in attempts at comparing distributions selected from varying economic political, and social environments.

More recent examples of this approach are Thurow's treatment of job competition and much of the literature on welfare program impacts. Most of these models are specific to a particular policy in that the distribution is not studied directly; instead, direct changes in income are calculated or analyzed. The very specificity of the approach limits, to some degree, the applicability of the models and the methodology, particularly when quantification of the institutional constraints is difficult.

EVALUATION OF THE DISTRIBUTION MEASURES (1, 2, 9, 17, 30, 34)

When methodologies are enumerated, usually some suggestions are made as to the "best" methodologies. One of the major problems with an evaluation of the income distribution methodologies is that the meaning of "inequality" is yet to be clearly established, so that the measures of "inequality" may be ambiguous. There are some criteria, however, which must be satisfied in order that the measures be consistent. These criteria are (1) impartiality

with respect to persons; (2) invariance with respect to numbers of persons; (3) invariance with respect to a uniform increase or decrease in the size of incomes; and (4) if two individual's incomes are changed while total income remains the same, the index must increase or decrease according to the absolute change between the two incomes. All the indices discussed in the paper meet these criteria. A ranking of the approaches depends upon the requirements of the research. It does appear to be reasonable, however, to conclude that the ambiguity of the formula measures suggest that the functional measures provide a much clearer definition of changes in income distribution.

Several researchers have attempted to establish criteria for judging the indices for a given set of criteria to determine the "best" of the measures. A ranking of indices or methodologies is performed according to the criteria.

Gastwirth has established a set of upper and lower bounds to the Gini coefficient, with which the functional forms can be evaluated according to the Gini coefficient which each produces. This is not a statistical test, however; it simply establishes the bounds on the Gini from an approach which does not assume a functional form of the distribution.

Finally, a graphical analysis can be performed. The Lorenz curve is graphed and the various indices and results from the various methodologies are drawn to determine which "best" fits the data given the research requirements. Clearly, these "tests" of appropriateness are wanting in rigor. As yet, statistical measures of appropriateness have not been found for a general case.

FUTURE PROBLEMS

Most of the future problems will be the "old" problems in a new form. As long as "inequality" remains inadequately defined, value judgments will necessarily enter into the selection of methodologies. As more information and research is done, the definitional problem will be less critical.

The data also present some obstacles to selection of the best methodology. While income data are available, wealth data are almost non-existent. The form of the data--that is, decrete quantiles--and the assumptions which are made when these quantiles are used--such as assuming the midpoint of a quantile is the income for all persons in that quantile--restrict the statistical power of at least the functional approximations. Further, detailed data for the nation are collected only once every ten years. National policies, such as natural resource policies, which affect changes in the whole economy and in most regions can not be evaluated with precision by using data which is so sparse. In addition, the comparability of the data is suspect, in that often the very definition of income changes from census to census. At a minimum, the practice of changing quantiles for which information is aggregated imposes a high cost to research efforts.

Finally, selecting the "best" from among all the methodologies is difficult. When distributions change--that is, the Lorenz curve shifts--often one methodology or functional form will not be the "best" fit for every distribution. There has been as yet no set of criteria established on which a choice can be made for all research. Thus, the consistency and comparability of results among research efforts are limited. At the same time, it should be expected that the testing of many more density functions and indices will continue. There is, and will be, a plethora of methodologies,

each applying to a limited set of research problems. It may well be that unless and until substantial progress is made on defining or approximating social welfare functions, choice of distributional methodologies will remain dependent on the specific researcher's problem. On the other hand, it does seem reasonable to suggest that the formula measures are generally inferior to the functional measures. One can hope that as information and theoretical advances occur, the ability to choose the appropriate mehtodology will be improved.

REFERENCES

1. Allingham, M. G. "Notes, Comments, and Letters to the Editor, the Measurement of Inequality." Journal of Economic Theory 5:163-169. 1972.

2. Atkinson, Anthony B. "On the Measurement of Inequality." Journal of Economic Theory 2: 244-263. 1970.

3. Becker, Gary S. and B. R. Whiswick. "Education and the Distribution of Earnings." American Economic Review 56(5):368-387. 1966.

4. Boulding, Kenneth E. and Martin Pfaff. Redistribution to the Rich and the Poor. Wadsworth Publishing Company, Belmont, California. 390 p.

5. Bronfenbrenner, Martin. Income Distribution Theory. Adline-Atherte Inc., New York. 487 p. 1971.

6. Budd, Edward C. Inequality and Poverty. W. W. Norton and Company, Inc., New York. 217 p. 1967.

7. Burgess, David F. "Tariffs and Income Distribution: Some Empirical Evidence for the U.S." Journal of Political Economics 84(1):17-45. 1976.

8. Champernowne, D. G. The Distribution of Income Between Persons. Cambridge University Press, Cambridge. 286 p. 1973.

9. Champernowne, D. G. "A Comparison of Measures of Inequality of Income Distribution." The Economic Journal 11:787-815. 1974.

10. Chipman, John S. "The Welfare Ranking of Pareto Distributions." Journal of Economic Theory 9:275-282. 1974.

11. Dasgupta, Partha, Sen. Amartya, and David Starrett. "Notes on the Measurement of Inequality." Journal of Economic Theory 6:180-187. 1973.

12. Elteto, O. and E. Frigges. "New Income Inequality Measures as Efficient Tools for Casual Analysis and Planning." Econometrica 36(2):383-396. 1968.

13. Friedman, M. "Choice, Chance, and the Distribution of Personal Income." Journal of Political Economy 61(4):277-290. 1953.

14. Gastwirth, Joseph L. "Notes and Comments: A General Definition of the Lorenz Curve." Econometrica 39(6):1037-1039. 1971.

15. Gastwirth, Joseph L. "The Estimation of the Lorenz Curve and Gini Index." The Review of Economics and Statistics LIV(3):306-316. 1972.

16. Gastwirth, Joseph L. "Large Sample Theory of Some Measures of Income Inequality." Econometrica 42(1):191-196. 1974.

17. Gastwirth, Joseph L. "The Estimation of a Family of Measures of Economic Inequality." Journal of Econometrics 3:61-70. 1975.

18. Hamada, Koichi. "A Simple Majority Rule on the Distribution of Income." Journal of Economic Theory 6:243-264. 1973.

19. Huffaver, G. C. "Environmental Quality, Income Distribution, and Factor Mobility: The Consequences of Local Income." Journal of Economic Issues VII(2):323-335. 1973.

20. Johnson, Harry G. and Peter Mieszkowski. "The Effect of Unionization on the Distribution of Income: A General Equilibrium Approach." The Quarterly Journal of Economics LXXXIV(11):539-600. 1974.

21. Kakwani, N. C. "A Note on the Efficient Estimation of the New Measures of Income Inequality." Econometrica 42(3):597-600. 1974.

22. Keith, John E., Ming C. Chen, Don C. Reading, and Robert Johnson. "An Empirical Analysis of Predictions of the Income Distribution Impacts of Water Quality Controls." ORS, Utah Water Research Laboratory Publication No. PRAAE 805-1. Logan, Utah; Utah State University. 1976.

23. Kondor, Yaakov. "An Old-New Measure of Income Inequality." Econometrica 39(6):1041-1042. 1971.

24. Levine, Daniel B. and Neil M. Singer. "The Mathematical Relation Between the Income Density Function and the Measurement of Income Inequality." Econometrica 38(2):324-330. 1970.

25. Martic, Ljubomir. "A Geometrical Note on New Income Inequality Measures." Econometrica 38(6):936-937. 1970.

26. Metcalf, Charles E. "The Size Distribution of Personal Income During the Business Cycle." The American Economic Review 59(9):657-668. 1969.

27. Minsky, Hyman P. "Effect of Shifts of Aggregate Demand Upon Income Distribution." American Journal of Agriculture Economics 50(2): 328-341.

28. Newberry, David. "A Theorem on the Measurement of Inequality." Journal of Economic Theory 2(9):264-266. 1970.

29. Newhouse, Joseph P. "A Simple Hypothesis of Income Distribution." U.S. Department of Commerce, Springfield, VA. 16 p. 1969.

30. Rothschield, Michael and Joseph E. Stiglitz. "Some Further Results on the Management of Inequality." Journal of Economic Theory 6:188-204. 1973.

31. Salem, A. B. A. and T. D. Mount. "A Convenient Descriptive Model of Income Distribution: The Gamma Density." Econometrica 42(6): 1115-1127. 1974.

32. Schultz, T. Paul. "Long Term Changes in Personal Income Distribution: Theoretical Approaches, Evidence and Explanations." MTIS No. P-4767 Washington, D.C.: NTIS. 1972.

33. Scoville, James C. Perspectives on Poverty and Income Distribution. D. C. Heath and Company, Lexington, Mass. 312 p. 1971.

34. Sen. Amartya. On Economic Inequality. W. W. Norton & Company, Inc., New York. 118 p. 1973.

35. Singh, S. K. and G. S. Maddala. "A Function for Size Distributions of Income." Econometrica 44:6. 1976.

36. Smith, James D. The Personal Distribution of Income and Wealth. Columbia University Press, New York. 568 p. 1976.

37. Thvrow, Lester C. "The Effects of Public Policy Instruments on Black and White Income Equalization." Redistribution to the Rich and the Poor. Belmont, California. Wadsworth Publishing. 1972.

38. Thvrow, Lester C. Generating Inequality. Basic Books, Inc. New York. 258 p. 1975.

39. Tinbergen, Jan. Income Distribution Analysis and Policies. American Elsevier Publishing Company, Inc., New York. 170 p. 1976.

40. U.S. Bureau of Census. Income Distribution of the United States. (Monograph prepared by Herman P. Miller.) U.S. Government Printing Office, Washington, D.C.

41. Williams, Gerald W. Impact of Federal Procurement on Personal Income Distribution. Unpublished Ph.D. dissertation. Washington University. 1968.